AMENDMENTS

In the Claims

The following is a marked-up version of the claims with the language that is underlined ("___") being added and the language that contains strikethrough ("~~—~~") being deleted:

1. (Currently Amended) A method comprising:

(A) receiving an email message from a simple mail transfer protocol (SMTP) server, the email message comprising displaying characters and non-displaying characters, the email message further comprising:

(A1) a 32-bit string indicative of the length of the email message;

(A2) a text body;

(A3) an SMTP email address;

(A4) a domain name corresponding to the SMTP email address;

(A5) an attachment;

(B) searching for the non-displaying characters in the email;

(C) removing the searched non-displaying characters;

(D) determining non-alphabetic displaying characters in the email;

(E) filtering the determined non-alphabetic displaying characters from the email;

(F) generating a phonetic equivalent for each word that includes only alphabetic displaying characters that has a phonetic equivalent;

~~(B)~~ (G) tokenizing the phonetic equivalents in the text body to generate tokens representative of words in the text;

~~(C)~~ (H) tokenizing the SMTP email address to generate a token representative of the SMTP email address;

~~(D)~~ (I) tokenizing the domain name to generate a token that is representative domain name;

~~(E)~~ (J) tokenizing the attachment to generate a token that is representative of the attachment, wherein tokenizing comprises:

~~(E1)~~ (J1) generating a 128-bit MD5 hash of the attachment;

~~(E2)~~ (J2) appending the 32-bit string to the generated MD5 hash to produce a 160-bit number; and

~~(E3)~~ (J3) UUencoding the 160-bit number to generate the token representative of the attachment;

~~(F)~~ (K) determining a probability value for each of the generated tokens;

~~(G)~~ (L) sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens;

~~(H)~~ (M) selecting ~~a~~ the predefined number of interesting tokens, the interesting tokens being the generated tokens having the greatest non-neutral probability values;

~~(I)~~ (N) performing a Bayesian analysis on the selected interesting tokens to generate a spam probability; and

~~(J)~~ (O) categorizing the email message as a function of the generated spam probability.

2. – 5. (Canceled)

6. (Currently Amended) A method comprising:

receiving an email message comprising a text body, an SMTP email address, an attachment, and a domain name corresponding to the SMTP email ~~address;~~ address, the text body including displaying characters and non-displaying characters;

searching for the non-displaying characters in the email;

removing the searched non-displaying characters;

tokenizing the SMTP email address to generate a token representative of the displaying characters of the SMTP email address;

tokenizing the attachment to generate a token that is representative of the attachment;

tokenizing the domain name to generate a token representative of the domain name;

determining a spam probability value from the generated tokens; and

sorting the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated tokens.

7. – 10. (Canceled)

11. (Previously Presented) The method of claim 6, wherein determining the spam probability comprises:

assigning a spam probability value to the token representative of the SMTP email address;

assigning a spam probability value to the token representative of the domain name; and

generating a Bayesian probability value using the spam probability values assigned to the tokens.

12. (Previously Presented) The method of claim 11, wherein determining the spam probability further comprises:

comparing the generated Bayesian probability value with a predefined threshold value.

13. (Previously Presented) The method of claim 12, wherein determining the spam probability further comprises:

categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold.

14. (Previously Presented) The method of claim 12, wherein determining the spam probability further comprises:

categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold.

15. (Canceled)

16. (Previously Presented) The method claim 6, wherein receiving the email message further comprises:

receiving an email message including a text body.

17. (Previously Presented) The method of claim 16, further comprising:

tokenizing the words in the text body to generate tokens representative of the words in the text body.

18. (Canceled)

19. (Previously Presented) The method of claim 17, wherein determining the spam probability comprises:

assigning a spam probability value to each of the tokens representative of the words in the text body;

assigning a spam probability value to the token representative of the attachment; and

generating a Bayesian probability value using the spam probability values assigned to the tokens.

20. (Previously Presented) The method of claim 19, wherein determining the spam probability further comprises:

comparing the generated Bayesian probability value with a predefined threshold value.

21. (Previously Presented) The method of claim 20, wherein determining the spam probability further comprises:

categorizing the email message as spam in response to the Bayesian probability value being greater than the predefined threshold.

22. (Previously Presented) The method of claim 20, wherein determining the spam probability further comprises:

categorizing the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold.

23. (Currently Amended) A system comprising:

email receive logic configured to receive an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an ~~address;~~ address, the email message further including displaying characters and non-displaying characters;

searching logic configured to search for the non-displaying characters in the email;

removing logic configured to remove the searched non-displaying characters;

tokenize logic configured to tokenize the SMTP email address to generate a token representative of the SMTP email address;

tokenize logic configured to tokenize the attachment to generate a token that is representative of the attachment;

tokenize logic configured to tokenize the domain name to generate a token representative of the domain name;

analysis logic configured to determine a spam probability value from the generated tokens; and

sorting logic configured to sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

24. (Currently Amended) A system comprising:

means for receiving an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an ~~attachment;~~ address, the email message further including displaying characters and non-displaying characters;

means for searching for the non-displaying characters in the email;

means for removing the searched non-displaying characters;

means for tokenizing the SMTP email address to generate a token representative of the SMTP email address;

means for tokenizing the attachment to generate a token that is representative of the attachment;

means for tokenizing the domain name to generate a token representative of the domain name;

means for determining a spam probability value from the generated tokens; and

means for sorting the generated tokens in accordance with the corresponding

determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

25. (Currently Amended) A computer-readable medium that includes a program that, when executed by a computer, performs at least the following:

receive an email message comprising an SMTP email address, a domain name corresponding to the SMTP email address, and an ~~attachment;~~ address, the email message further including displaying characters and non-displaying characters;

search for the non-displaying characters in the email;

remove the searched non-displaying characters;

tokenize the SMTP email address to generate a token representative of the SMTP email address;

tokenize the attachment to generate a token that is representative of the attachment;

tokenize the domain name to generate a token representative of the domain name;

determine a spam probability value from the generated tokens; and

sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

26. (Previously Presented) The computer-readable medium of claim 25, the program further causing the computer to perform at least the following:

assign a spam probability value to the token representative of the SMTP email address;

assign a spam probability value to the token representative of the domain name; and

generate a Bayesian probability value using the spam probability values assigned to the tokens.

27. (Previously Presented) The computer-readable medium of claim 26, the program further causing the computer to perform at least the following:

compare the generated Bayesian probability value with a predefined threshold value.

28. (Previously Presented) The computer-readable medium of claim 27, the program further causing the computer to perform at least the following:

categorize the email message as spam in response to the Bayesian probability value being greater than the predefined threshold.

29. (Previously Presented) The computer-readable medium of claim 27, the program further causing the computer to perform at least the following:

categorize the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold.

30. (Currently Amended) A system comprising:

a memory component that stores at least the following:

email receive logic configured to receive an email message comprising an ~~attachment;~~ address, the email message further including displaying characters and non-displaying characters;

search logic configured to search for the non-displaying characters in the email;

remove logic configured to remove the searched non-displaying characters;

tokenize logic configured to tokenize the attachment to generate a token representative of the attachment;

analysis logic configured to determine a spam probability value from the generated token; and

sort logic configured to sort the generated tokens in accordance with the corresponding spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

31. (Currently Amended) A system comprising:

means for receiving an email message comprising an ~~attachment;~~ address, the email message further including displaying characters and non-displaying characters;

means for searching for the non-displaying characters in the email;

means for removing the searched non-displaying characters;

means for tokenizing the attachment to generate a token representative of the attachment;

means for determining a spam probability value from the generated token; and

means for sorting the generated tokens in accordance with the corresponding

determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

32. (Currently Amended) A computer-readable medium that includes a program that, when executed by a computer, performs at least the following:

receive an email message comprising an ~~attachment;~~ address, the email message further including displaying characters and non-displaying characters;

search for the non-displaying characters in the email;

remove the searched non-displaying characters;

tokenize the attachment to generate a token representative of the attachment;

determine a spam probability value from the generated token; and

sort the generated tokens in accordance with the corresponding determined spam probability value to determine a predefined number of interesting tokens, the predefined number of interesting tokens being a subset of the generated ~~tokens.~~ tokens, wherein only displaying characters are tokenized.

33. (Previously Presented) The computer-readable medium of claim 32, the program further causing the computer to perform at least the following:

receive an email message having a text body.

34. (Previously Presented) The computer-readable medium of claim 33, the program further causing the computer to perform at least the following:

tokenize the words in the text body to generate tokens representative of the words in the text body.

35. (Previously Presented) The computer-readable medium of claim 34,

assign a spam probability value to each of the tokens representative of the words in the text body;

assign a spam probability value to the token representative of the attachment; and

generate a Bayesian probability value using the spam probability values assigned to the tokens.

36. (Previously Presented) The computer-readable medium of claim 35, the program further causing the computer to perform at least the following:

compare the generated Bayesian probability value with a predefined threshold value.

37. (Previously Presented) The computer-readable medium of claim 36, the program further causing the computer to perform at least the following:

categorize the email message as spam in response to the Bayesian probability value being greater than the predefined threshold.

38. (Previously Presented) The computer-readable medium of claim 36, the program further causing the computer to perform at least the following:

categorize the email message as non-spam in response to the Bayesian probability value being not greater than the predefined threshold.

39. (Previously Presented) The method of claim 1, wherein the email is received at a computing device.